

# README File for 2010 Census American Samoa (AS) Summary File Delivered via FTP

***Note:** Users processing these FTP files in a Windows environment should read carefully the File Information section of this document.*

*This revised file (Revision 1) intentionally excludes tables: HBG72-78, HCT17-19, and PBG7 & PBG9. The table segmentation has been updated to reflect this change. For additional information, please refer to the Technical Documentation: <http://www.census.gov/prod/cen2010/doc/sfas.pdf>*

## **Contents**

*About the FTP Application*

*The FTP Directory*

*File Information*

*Segmented Data*

*File Record Layout*

## **About the FTP Application**

**This FTP (File Transfer Protocol) application is intended for experienced users of census data, zip files, and spreadsheet/database software.** It provides quick access for data users such as State Data Centers and news media needing to begin their analysis immediately upon data release. Due to the potentially large size of the files, the FTP user should have a fast file transfer capability.

## **The FTP Directory**

The directory is [http://www2.census.gov/census\\_2010/09-Island\\_Areas\\_IASF/American\\_Samoa/](http://www2.census.gov/census_2010/09-Island_Areas_IASF/American_Samoa/). When the American Samoa Summary File data set is added to the area's directory, it will contain a single zip package with the geoheader file, the 31 data segments and a packing list which provides information about the file's creation and size. See below for more information on Segmented Data.

Users of the FTP application need to unzip the package after downloading, then import the data into the spreadsheet/database software of their choice for data analysis and table presentation. *We are unable to provide one-on-one support for applications of the data to specific spreadsheets or data base software.*

## **File Information**

Once the package is unzipped, the files are in flat ASCII format. The geographic header file contains fixed fields while the data files are in comma-delimited format. These are text files however the file extension is **not** '.txt'. The user will need to rename the files with a .txt extension for import into some software packages, e.g. Microsoft Access. These files have been constructed in a LINUX environment. They use an ASCII linefeed, chr(10), to indicate a new record.

For successful use with many programs running in a Windows environment, these files need to be modified to use the ASCII carriage return/linefeed sequence, chr(13) + chr(10) as a record terminator. This is an easy step in the UnZIP process using any UnZIP software which offers the conversion option. WinZip for Windows, version 14.5 has been tested.

The resulting file will meet the ANSI MS-DOS/Windows standard used by Access 2003 and Access 2007 and other MS Windows-based programs. If the data are being processed in a LINUX environment, they can be unzipped using any standard Linux ZIP/UnZIP package.

### **Note to Users of Microsoft Access:**

Due to the FieldSize property limitations within Microsoft Access, modifications to field types are required when importing the Geographic Header Record file and File01 through File31:

- Fields classified as numeric (N) should be imported as long integers unless the field has been described as having decimals. These fields should be imported as double.
- LOGRECNO, AREALAND and AREAWATR should be imported as text.
- Failure to make these changes may result in missing data upon import.

## **Segmented Data**

The data in the 2010 Census American Samoa Summary File and other 2010 Census summary files are segmented. This is done to manage the volume of data and to facilitate exporting into spreadsheet or database software. The data and the corresponding geographic information for an individual state/island area are known as the file set. Because of the large size of the tables, the file set will be broken into 33 files: a geographic header record file, 31 data segment files and a packing list. To get a complete set of the American Samoa Summary File data, users must download the geographic header file and all the data file segments in the package. See Table 1 below for segmentation details.

**Table 1. American Samoa File / Table Segmentation**

Data file segment number	Number of data cells	Starting matrix number	Ending matrix number
1	232	P1	P13
2	226	P14	P27
3	113	P28	P36
4	198	PBG1	PBG21
5	230	PBG22	PBG37
6	229	PBG38	PBG78
7	230	PBG79	PBG125
8	66	PBG126	PBG139
9	209	PCT1	PCT1
10	106	PCT2	PCT8
11	231	PCT9	PCT13
12	231	PCT14	PCT22
13	219	PCT23	PCT29
14	234	PCT30	PCT35
15	218	PCT36	PCT43
16	234	PCT44	PCT45
17	193	PCT46	PCT51
18	168	PCT52	PCT57
19	123	PCT58	PCT62
20	202	PCT63	PCT66
21	187	PCT67	PCT72
22	205	PCT73	PCT75
23	231	PCT76	PCT78
24	122	PCT79	PCT86
25	102	H1	H14
26	235	HBG1	HBG26
27	235	HBG27	HBG61
28	164	HBG62	HBG109
29	10	HBG110	HBG111
30	221	HCT1	HCT7
31	220	HCT8	HCT16

The explanation below for linking the two data files requires specific location information about the geographic header. These are located in Chapter 2-How to use this Product, of the Technical Documentation <http://www.census.gov/prod/cen2010/doc/sfas.pdf>.

A unique logical record number (LOGRECNO) in the geographic header is assigned to all the files for a specific geographic entity, linking all the records for that entity across file segments. The geographic header **record layout** is identical across all electronic data products from the 2010

Census. However, the content is product specific. Some header fields that appear in both file types (geographic header and fileXX) are not used. For example, the characteristic iteration (CHARITER) field is filled with a valid code in the 2010 Census Summary File 2 data product but in the 2010 Census American Samoa Summary File, it is always coded as 000.

### **File Record Layout**

For a layout of the data table, see <http://www.census.gov/prod/cen2010/doc/sfas.pdf>, select Chapter 6, *Data Dictionary*.